Young children's reasoning about the order of past events

McCormack, T., & Hoerl, C. (2007). Young children's reasoning about the order of past events. *Journal of Experimental Child Psychology*, *98*(3), 168-183. https://doi.org/10.1016/j.jecp.2007.06.001

Published in:
Journal of Experimental Child Psychology

Document Version:
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Journal of Experimental Child Psychology 98 (2007) 168–183

Journal of Experimental Child Psychology

www.elsevier.com/locate/jecp

Young children's reasoning about the order of past events ☆

Teresa McCormack [a,*], Christoph Hoerl [b]

[a] *School of Psychology, Queen's University, Belfast BT7 1NN, UK*
[b] *Department of Philosophy, University of Warwick, Coventry, CV4 7AL., UK*

Received 5 February 2007; revised 13 June 2007
Available online 3 August 2007

Abstract

Four studies are reported that employed an object location task to assess temporal–causal reasoning. In Experiments 1–3, successfully locating the object required a retrospective consideration of the order in which two events had occurred. In Experiment 1, 5- but not 4-year-olds were successful; 4-year-olds also failed to perform at above-chance levels in modified versions of the task in Experiments 2 and 3. However, in Experiment 4, 3-year-olds were successful when they were able to see the object being placed first in one location and then in the other, rather than having to consider retrospectively the sequence in which two events had happened. The results suggest that reasoning about the causal significance of the temporal order of events may not be fully developed before 5 years.

Keywords: Time; Reasoning; Causation

There are good reasons to believe that by 4 years children are adept at dealing with temporal order information. Children of this age can verbally describe both routine and novel event sequences in the correct order (see Fivush & Hudson, 1990; Nelson, 1986), and even much younger children have expectations about the order in which events in a familiar sequence occur, and can reproduce novel observed sequences of events in the correct order (Bauer, Wenner, Dropik, & Wewerka, 2000; Carver & Bauer, 2001). In fact, basic forms of

☆ This research was supported by an interdisciplinary grant from the Arts and Humanities Research Council, UK. We are grateful to the nursery and primary schools that assisted us with the studies, and to Helena McCormack who tested the children. We also thank Fred Maddalena for constructing the materials.
* Corresponding author. Fax: +44 02890 664144.
E-mail address: t.mccormack@qub.ac.uk (T. McCormack).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2007.06.001

perception and memory for temporal order have been demonstrated in infants of only a few months' age (Lewkowitz, 2004; Mandel, Nelson, & Jusczyk, 1996). Four-year-olds' judgments about how to complete sequences of pictures representing familiar types of causal changes also demonstrate that they are sensitive to the order in which the events in the sequence occur (Das Gupta & Bryant, 1989; Gelman, Bullock, & Meck, 1980), and they can easily distinguish object transformations that occur in the correct forwards direction from those that are shown in an anomalous backwards direction (Friedman, 2003a). The considerable body of research conducted by William Friedman (see Friedman, 2003b, for review) on young children's temporal cognition has demonstrated that 4-year-olds can make judgments about the relative order in which significant events (e.g., their birthday and Christmas) occurred in their past (Friedman, Gardener, & Zubin, 1995; Friedman & Kemp, 1998), and can also do so for at least some pairs of novel events they have encountered on a single occasion (Friedman, 1991; though see Friedman & Lyon, 2005). Four-year-olds can also correctly order representations of sequences of familiar events, such as the main events of a typical day or a visit to the supermarket (Fivush & Mandler, 1985; Friedman, 1977; Friedman & Brudos, 1988).

All of this evidence suggests that children of this age can encode and remember temporal order information, and make use of it insofar as they can judge whether event sequences are correctly or incorrectly ordered. However, it leaves open the extent to which they can incorporate information about temporal order into further reasoning. It has recently been suggested that it is not until around 5 years that children can fully understand and thus fully make use of information about the temporal order in which events have occurred (McCormack & Hoerl, 2005; Povinelli, Landry, Theall, Clark, & Castille, 1999). Specifically, Povinelli et al. (1999) have suggested that although young children may indeed possess temporally ordered representations of the world, their reasoning is constrained because they do not fully grasp the underpinnings of and the significance of such chronological organization. They point out that as mature thinkers we are aware that the chronological organization of events reflects the causal relationships that exist between different events. Povinelli et al.'s suggestion is that an adult-like concept of time involves not only the ability to represent the temporal order of events, but also an appreciation of the causal significance of that temporal order. In other words, a mature understanding of time involves grasping that the passage of time leads to the unfolding of "a successive series of causally interdependent states of the world" (Povinelli et al., 1999, p. 1427).

The idea that a mature or objective conception of time involves a certain type of causal understanding is one that has appeared previously in the philosophical literature (notably, Campbell, 1994; see also Martin, 2001). If this idea is correct, then providing a complete account of the development of temporal concepts would involve considering not just the age at which children can make judgments about or reconstruct the temporal order of events, but evidence that they understand the causal significance of temporal order. Without such a grasp, errors may occur in inferential reasoning contexts. For example, because later events in an event sequence may change or obliterate the effects of events that occurred earlier in the sequence, the overall outcome of a sequence of events often depends not just on what events have happened but also on the temporal order in which they happened. As a result, failure to consider the order of occurrence of events in a sequence when reasoning about the causal consequences of those events can lead to errors. Indeed, Povinelli et al.'s (1999) and McCormack and Hoerl (2005) findings can be interpreted as evidence that young children are prone to such errors.

In Povinelli et al.'s (1999) study, children took part in two games, one after another with a short delay between them. Behind the child's back and without the child noticing, an experimenter placed a toy in a box behind the child during the playing of Game 1, and then moved it and placed it in another box during the playing of Game 2. Children were subsequently shown videotapes of themselves playing the games, with the experimenter's actions now visible in the background. They were then asked to judge the toy's current location. The crucial manipulation in the study was that the videotapes of the games were not necessarily shown in the order in which the games had been played: half of the time children were shown where the experimenter hid the toy during Game 1 and then where it was moved during Game 2, but half of the time they were shown first what happened during Game 2 and then what happened during Game 1. That is, the order in which children found out about each hiding event was decoupled from the order in which the hiding events had occurred. As a consequence, in Povinelli et al.'s study, children did not just have to remember the order in which the games had been played, but also had to use this order information to work out where the toy must now be (i.e., in the location it was placed in during Game 2). It was found that children below 5 years were unable to use the order information in this way, even when cued to consider it. Povinelli et al. described this failure as due to a lack of understanding of the causal significance of the temporal order in which the events took place.

McCormack and Hoerl (2005) used a quite different task, but obtained similar results. In their task, children were initially introduced to two dolls, Sally and Katy, who always performed actions in a certain order: they learned that Sally always went first, and then Katy always went next. Children were then shown a novel piece of apparatus—a large yellow box with two differently colored buttons—and learned how it worked. Pressing one of the buttons caused a toy car to drop down one chute and appear on a shelf in a transparent window, whereas pressing the other button caused a marble to drop down another chute and appear on the shelf. The window only ever contained one toy at a time, as the box was mechanically constructed such that whatever toy was already in the window dropped out of sight into a drawer below before a new toy appeared. Children had to learn which button yielded which toy. At test, a screen was then put in front of the box and, out of view of the child, Sally and Katy pressed one button each. The nature of the box was such that which object was left in the window after these button-presses depended on the order in which the buttons were pressed. Finally, the box was uncovered again.

In one version of the task, after the screen was removed, children could see each doll standing next to the button that she had pressed, but the window in the box was left occluded. In this version, children had to infer which toy was inside the window. In other words, given information from which they could determine the order in which two earlier events had occurred (the pressing of the buttons by the dolls), children had to infer the current state of the world (the object in the window). Four-year-olds consistently failed this task, whereas 5-year-olds passed, at least when cued to attend to the order information. McCormack and Hoerl (2005) described their task as assessing temporal-causal reasoning, which they defined as reasoning in which "the causal implications of the temporal order of two or more events must be considered in one's reasoning" (p. 55). In both their task and that of Povinelli et al., inferences about the current state of the world will only be correct if the order in which the two relevant events have occurred is considered; thus both tasks can be described as temporal–causal-reasoning tasks.

The issue addressed in the current study is whether the paradigms used by Povinelli et al. (1999) and McCormack and Hoerl (2005) have led to an underestimation of young children's abilities. Importantly, Povinelli et al.'s task required children to understand something about the representational medium of videotape—in particular that video clips of events that are novel to them in fact depict real events that occurred in the recent past, and that the order in which two events are presented on video is not necessarily determined by the order in which those events actually happened. Although Povinelli et al. attempted to ensure that children understood the nature of video, it seems sensible to assess whether children still have difficulties with temporal-causal reasoning when the task does not involve this representational medium, aspects of whose functioning they may still be unfamiliar with.

McCormack and Hoerl's (2005) task was an attempt at such an assessment. Although their findings closely resembled those of Povinelli et al. (1999) using a quite different task, it is still necessary to consider the additional demands that their task may have placed on young children. Arguably, the novelty of their apparatus, and the demands the task placed on working/long-term memory may have led to an underestimation of 4-year-olds' abilities. In the task, children first had to learn which button press gave which object. At test they then had to retrieve this information and put it together with information about the order in which the buttons had been pressed by the dolls, which in itself involved retrieving information about the order in which the dolls always acted. Retrieving these pieces of information and then combining them may have proved beyond the working memory capabilities of young children. Indeed, McCormack and Hoerl (2005) distinguished between two quite different explanations of 4-year-olds' failure: that children of this age genuinely do not seem to grasp the significance of temporal order information, or that the processing demands of the task (in particular, the working-memory demands) are beyond their capabilities. They admitted that their findings alone did not allow them to decide between these two possibilities.

Thus, the aim of the present study was to examine 4-year-olds' abilities using a task that had a minimal working-memory load. It was also important to use a task that did not require competence with a potentially unfamiliar representational medium such as videotape, as required by Povinelli et al.'s (1999) study. As in McCormack and Hoerl's (2005) task, the task involved two dolls that always acted in a particular order. However, there were a number of key differences between their task and the one used in the present study. First, and most importantly, the consequences of the two dolls' actions were essentially given to children in the display at test. This contrasts with McCormack and Hoerl's task, which required retrieval of novel pairings from long-term memory that had to be learned in a pre-test training session. In the present study, children simply needed to combine the information that was clearly given to them in the test display with information about the order in which the dolls had acted. Second, control questions were always asked immediately before the test question that would have acted as a cue to retrieve such order information. These control questions ensured that children had in mind the order the dolls carried out the relevant actions just before they had to answer the test question. Both of these procedural adjustments ensured that the memory demands of the task were kept to a minimum.

There were three additional differences. Unlike in McCormack and Hoerl's (2005) task, the task did not involve a completely novel piece of mechanical apparatus; rather the relevant events took place in a prototypical doll's house. Furthermore, the dolls' actions were set in the context of a simple narrative to engage children's attention, whereas there was no such narrative underpinning the events in McCormack and Hoerl's task. The last difference concerns the events that necessarily took place out of sight of the child during the procedure. In both McCormack and Hoerl's and Povinelli et al.'s (1999) task, there was no rationale provided to children for why the crucial events had occurred out of sight of the child. Thus, young children may not have felt confident that there was a correct answer that could be logically inferred from the evidence (e.g., because they may have assumed that trickery had been carried out). In the present study, the fact that the dolls' actions took place out of sight made sense within the task context. Finally, although the task differed from those used in previous studies in the ways that have been outlined, it was in one important respect similar: children encountered just a single example of the relevant event sequence. This was to ensure that their answers were based on reasoning about event order information, rather than based on accrued knowledge about the typical outcomes of particular sequences. We return to this issue in the General Discussion.

The findings from Povinelli et al.'s (1999) and McCormack and Hoerl's (2005) studies are striking because the pre-existing research literature might lead one to predict that 3- or 4-year-olds should pass such a task. As we have pointed out, there is considerable evidence to suggest that even very young children are competent at representing and remembering temporal order information (reviewed in Friedman, 2003b). Furthermore, their competence in a number of domains involving temporal and/or causal cognition has been well-demonstrated: existing research suggests that 3- or 4-year-olds have some planning abilities (e.g., Atance & O'Neill, 2001; Hudson, Sosa, & Shapiro, 1997), can make at least some simple future hypothetical and counterfactual judgments (e.g., Beck, Robinson, Carroll, & Apperly, 2006), can use temporally-structured scripts to guide their behavior (Nelson, 1986, 1996), and seem to systematically base their judgments of the causal powers of objects on patterns of evidence (e.g., Gopnik et al., 2004). Arguably, this body of research might lead us to predict that 4-year-olds should be successfully on the simplified task used in the present study.

Experiment 1

Methods

Participants

Fifty-six children took part in the study, 29 4-year-olds ($M = 53$ months; $Range = 48$–57 months; 16 boys and 13 girls) and 27 5-year-olds ($M = 65$ months; $Range = 61$–71 months; 9 boys and 18 girls). Children were recruited from nurseries and schools local to the university of the first author. They were tested individually in a quiet corner of their schools or nurseries, and each child received a sticker of his or her choice for taking part.

Materials

A purpose-built two-storey wooden dolls' house 55 cm wide, 38 cm tall, and 14 cm depth was used. There were four equally-sized rooms in the house (kitchen, living room, bedroom, and bathroom), with each room containing wooden furniture appropriate to the function of the room. In the bathroom there were two colored cupboards, each 12 cm tall, 5 cm wide and 3 cm in depth, one blue and one red. There was also one miniature toy hairbrush 2.5 cm long. The house was constructed with gaps in the front and back walls of each room; this allowed both the experimenter and the child to see into the rooms, with the experimenter seated behind the house facing the child, who sat in front of the house. There was a sliding door attached to the bathroom that the experimenter could close fully, preventing the child but not the experimenter from seeing into the room. Two wooden clothed dolls were used in the study, with different heights (11 and 9 cm), hair color, and clothing. At the test phase, two life-size laminated photographs of the bathroom cupboards were used.

Procedure

Children were shown the dolls' house and their attention was drawn to the bathroom. They were then shown the two photographs of the cupboards and asked to point to the cupboard in the house shown in each picture. No child had difficulty with this. Children were then introduced to the two doll characters, John and Peter, and told that they had been playing in the garden and had got a bit messy, and that they were now going into the house. They were told that the dolls were going to wash their hands in the kitchen sink, and as they did so the experimenter drew the child's attention to the order in which the dolls acted, and told the child that the dolls always acted in this order. The dolls then carried out two further activities (playing with a toy train in the living room and going on the rocking horse in the bedroom). The experimenter provided a similar narration to each activity that was similar to that used in the kitchen and emphasized the order in which the dolls always acted. The children were then told that the dolls were going to go into the bathroom to brush their hair which had got messy when they were playing outside. The hairbrush, which was sitting by the bathroom sink, was pointed out to them, as were the two cupboards. When the two dolls went into the bathroom, the experimenter closed the door. The house was constructed so that the experimenter still had access to the room while the door was closed, and the experimenter said "You can't see John right now, but he goes first and gets the hairbrush and now he is brushing his hair. Now he puts the hairbrush in one of the cupboards. Peter goes last. You can't see him now, but he gets the hairbrush out and now he is brushing his hair. Now he puts the hairbrush into the other cupboard," with the cupboard that each doll put the hairbrush in counterbalanced across trials. After this, the door to the room was opened and the dolls emerged with tidy hair.

The test phase immediately followed the emergence of the dolls from the bathroom, with dolls placed side-by-side in front of the child. Children were initially asked two control questions about the order of the dolls' actions, in counterbalanced order: "Can you remember which doll brushed his hair first?" and "Can you remember which doll brushed his hair last?". All children answered these questions correctly. The experimenter then reminded the child that the hairbrush was inside one of the cupboards, and said that she was going to put each doll beside a picture of the cupboard that he had put the hairbrush into. The photographs were placed beside the dolls, initially face-down, with left–right location of the cupboard that contained the hairbrush counterbalanced across trials. The experimenter turned over both photographs simultaneously, and then asked the child: "So, where do you think the brush is right now?", with either a verbal or a pointing response deemed acceptable. Regardless of whether or not the child was correct, the experimenter then thanked the child for playing the game well, and the child was given a sticker.

Results and discussion

Table 1 shows the proportion of children who selected the correct cupboard. The majority of children in each age-group selected the correct cupboard as the hairbrush's current location. However, although the 5-year-olds did choose the correct location significantly

Table 1
Percentage of children who passed the tasks in Experiments 1–3

		Correct %
Experiment 1	4-year-olds ($N = 29$)	62
	5-year-olds ($N = 27$)	81*
Experiment 2	4-year-olds ($N = 28$)	54
Experiment 3	4-year-olds ($N = 26$)	50

* $p < .05$.

more frequently than would have been expected by chance (binomial test, $p < .05$), as a group the 4-year-olds did not perform significantly above chance (binomial test, $p > .05$).

These findings are compatible with those of McCormack and Hoerl (2005), in that 5- but not 4-year-olds were able to use information about the order in which two events had happened to make a correct inference about a current state of the world. Unlike in McCormack and Hoerl's task, children did not have to retrieve information from long-term memory about the impact of each doll's action. The impact that each doll's action had in this task was on the location of the hairbrush, and the location of this object following each doll's action was directly given to participants in the photographs placed in front of them. These photographs were present while the test question was being asked. The fact that children did not have to retrieve such information from long-term memory to combine it with information about the order in which the dolls had acted makes it difficult to explain 4-year-olds' failure of the task in terms of a working memory problem.

Passing the task depended crucially on remembering and using information about the order in which the dolls had acted. It would seem difficult to explain 4-year-olds' performance in terms of a failure to remember such information, since all children answered control questions assessing this knowledge immediately before the test question was administered. Both Povinelli et al. (1999) and McCormack and Hoerl (2005) tried to boost younger children's performance by explicitly drawing their attention to the order of the relevant events (albeit without success). Asking children the control questions immediately before the test question should have served to draw their attention to event order. However, in Experiment 2, we aimed to make the event order even more salient by actually letting children see that the dolls acted in a particular order. In this version of the task, children observed the dolls carrying out the relevant actions one at a time, but could not actually infer the outcome of their actions before the crucial information was provided at test. This allowed children to directly encode the order in which the dolls had acted, and this alteration, it was hoped, should have made the order of the doll's actions more salient to them.

Experiment 2

Methods

Participants

Twenty-eight 4-year-olds ($M = 55$ months; $Range = 50–59$ months; 12 boys and 16 girls) took part in the study, none of whom had taken part in Experiment 1. They were recruited and tested in a manner identical to that of Experiment 1.

Materials

These were the same materials as were used in Experiment 1.

Procedure

The procedure was identical to that used in Experiment 1, with a single alteration: rather than the dolls both going into the bathroom at the same time, they went in one at a time. At the appropriate point in the procedure, the experimenter explained to children, as in Experiment 1, that the dolls were going to go into the bathroom to brush their hair. Then the first doll (John) went into the bathroom on his own, and the door was closed behind him. His unseen actions were described to the child as before (i.e., "Now he is brushing his hair. Now he puts the hairbrush in one of the cupboards..."), and on completing these actions, the door was opened and the doll emerged with tidy hair. When the door was opened, children could not see where the hairbrush was, as it was at that point already inside one of the cupboards. Following John's emergence, Peter went into the bathroom and the same procedure was repeated. The test phase began immediately after Peter emerged from the bathroom, and was identical to that used in Experiment 1.

Results and discussion

As can be seen from Table 1, this alteration in procedure did not improve performance. Children's performance did not differ significantly from chance levels (binomial test, $p > .05$). Thus, although the order in which the dolls acted was actually visible to children as the procedure was conducted, and thus could easily be encoded by them, 4-year-olds were nevertheless unable to use such information to answer the test question.

These results suggest that 4-year-olds' difficulties really do lie with realizing the significance of the order information that is available to them and using it to answer the test question. However, we wished to rule out the possibility that children's difficulties lie with the use of photographs of the object's location during the test phase. We initially decided to use photographs because they provided a way of informing children which cupboard each doll had placed the hairbrush in simultaneously for both dolls. Informing them of this verbally would have necessitated providing the information in a certain order (i.e., first about the location in which one doll had placed the hairbrush and then about the location in which the other doll had placed it), which would have raised the possibility of children simply updating their representation of the hairbrush's location sequentially as they found out each piece of information. This would have led to the wrong answer if the order in which they were told about each location visit differed from the order in which the visits had occurred. Although the use of photographs, which could be revealed simultaneously, ruled out this possibility, perhaps children in some way had difficulty understanding the significance of the photographs.

In our pre-test check, we ensured all children could match the photographs to the actual cupboards, and indeed previous research would indicate that 4-year-olds should have no difficulties with understanding the correspondences between photographs and real objects (see Liben, 2003). Nevertheless, given that we wish to argue that 4-year-olds' problems are with making the appropriate inference, it is important to rule out all explanations of their failure that are theoretically uninteresting. Thus, in the third experiment, rather than using photographs, each doll was placed immediately adjacent to the cupboard that he had put the hairbrush in. If children's difficulties on the previous tasks were with the photographs, they would be expected to be successful on this task.

Experiment 3

Methods

Participants

Twenty-six 4-year-olds ($M = 51$ months, $Range = 48–54$ months; 13 girls and 13 boys) completed the task, none of whom had taken part in the first two experiments. They were recruited and tested in a manner identical to that of Experiment 1.

Materials

These were identical to those used in the first two experiments, except that the photographs were not used.

Procedure

The task followed the procedure used in Experiment 1, up until the point at which the bathroom door was opened after the dolls had brushed their hair. In this experiment, the door was opened to reveal each of the dolls standing beside one of the cupboards (the specific cupboard that each doll stood beside was counterbalanced). Participants were then told that each doll was standing beside the cupboard that he had placed the hairbrush in, and then were asked the two control questions, followed by the test question.

Results and discussion

As can be seen in Table 1, children performed on this task at chance levels, with exactly half of the group choosing each location. This result rules out an explanation of 4-year-olds' failure on the first two experiments in terms of their difficulties using the photographs as depictions of the locations in which the hairbrush had been placed. It provides further support for the hypothesis that children of this age have genuine difficulties making use of temporal order information.

Experiment 4

In our last experiment, we wished to rule out any possibility that there is something about the test procedure itself which may distract or confuse children. Perhaps asking the child the two control questions rather than acting as a cue to consider temporal order information may somehow have confused them and led to difficulties focusing on the test question. The version of the task used in Experiment 4 followed the procedure used in Experiment 1, except that we actually allowed children to see the hairbrush being placed in each cupboard. The test phase then proceeded as in Experiment 1. Our rationale here was that if there was something about the procedure during the test phase that was distracting or confusing, we might expect children to make errors even though they had actually seen the hairbrush being placed first in one cupboard and then in the next. In other words, even though in this version of the task children could provide a correct answer simply through sequentially updating their representation of the hairbrush's location, if the procedure used at test is itself problematic we might expect to see poor performance. On the other hand, if children's problems are with making the appropriate inference, we would expect high levels of performance when the correct answer can be given by simply remembering where the hairbrush was last placed.

In this experiment, younger children (3-year-olds) were tested. This was because we wished to assess whether the findings from the present task would provide further evidence for the distinction between temporal-causal reasoning and the much more developmentally primitive temporal updating (McCormack & Hoerl, 2005). In temporal updating, children simply change their model of the world (e.g., where certain objects are located) when they receive information about changes in the world. Children may sequentially update a model in response to observing or inferring such changes, or perhaps due to being told about them. Crucially, children's models change over time simply as a result of them receiving information about different events in a certain order, rather than children themselves reasoning about the order in which those events happened. The altered version of the task used in Experiment 4 could be passed by simple updating since children actually viewed the hairbrush being placed in first one cupboard and then another, whereas this was not the case in the previous experiments. Given that McCormack and Hoerl (2005) argued that temporal updating is developmentally primitive, we might expect even 3-year-olds to find the task trivially easy, despite 4-year-olds' consistent difficulties on the task in Experiments 1–3.

Methods

Participants

Seventeen 3-year-olds ($M = 42$ months; *Range* = 34–47 months; 6 boys and 11 girls) took part in the experiment. These were recruited through a nursery school local to the first author's university. They were tested in their nursery, and received stickers for taking part in the task.

Materials

These were identical to those used in Experiment 1.

Procedure

This was identical to that used in Experiment 1 with a single modification: the bathroom door was left open at all times during the task. Small alterations were made to the experimenter's narration to make it compatible with this changed procedure (e.g., the experimenter did not say "You can't see X right now, but...."). The test phase was identical to that used in Experiment 1.

Results and discussion

All of the children except one chose the correct location; thus performance was significantly above chance on this task (binomial test, $p < .05$). This result suggests that when the task can be passed simply by observing and remembering the consequences of the dolls' actions, it is trivially easy even for 3-year-olds. This finding indicates that 4-year-olds' problems on the task seem to be with making the relevant temporal–causal inference, and that the procedure at test (e.g., asking the two control questions before the test question) itself does not unnecessarily confuse or distract children.

General discussion

The findings of Experiments 1–3 are entirely consistent with those of McCormack and Hoerl (2005) and Povinelli et al. (1999), in that 4-year-olds had difficulties using information about the order in which two events had occurred to reach a conclusion about the current state of the world. This ability has been labeled temporal–causal reasoning, since it requires a grasp of the causal significance of temporal order information. Four-year-olds had difficulties even though they correctly answered control questions about the order in which the dolls had acted that were asked immediately before test—questions which might have been expected to cue children to the importance of considering the temporal order of the dolls' actions. Furthermore, allowing children to actually observe the order in which the dolls acted (but not their actions; Experiment 2), did not improve performance. Experiments 1 and 2 required that children grasp the correspondence between life-size photographs depicting two cupboards and the two cupboards in the dolls' house, one of which they had to choose as the location of the target object. Although there was no reason to believe that children had difficulty with this, Experiment 3 used a version of the task in which photographs were not used, with the dolls simply placed beside the locations that each had placed the object in. This did not improve performance, suggesting that the task taps the ability to make a basic type of inference that young children have difficulty with.

In Experiment 4, an identical procedure was used, except that children were actually able to see the relevant actions of the dolls. This allowed children to simply sequentially update their representation of the object's location as they saw it being moved. Under such circumstances, the task proved easy even for 3-year-olds. This suggests that 4-year-olds' difficulties lie with making a retrospective inference about events which they did not observe at the time of occurrence. The finding provides further support for the distinction between temporal–causal reasoning and a much more developmentally primitive ability, temporal updating (McCormack & Hoerl, 2005). Performance in Experiment 4 could have been based on such updating, and indeed even 3-year-olds had no difficulties in such a task. The key difference between the task in Experiment 4 and the tasks in the other experiments was that it did not actually require the child to consider the order in which the events had happened in reaching a conclusion: it was sufficient just to retrieve a representation of the hairbrush's location that had been appropriately updated on observing its previous movements. Thus, the task in Experiment 4 did not require temporal–causal reasoning as we have described it.

We now turn to considering how our research findings relate to those of studies examining other aspects of temporal and causal cognition. Historically, much research on the development of causal cognition has focused on children's ability to select, among a number of candidates, the cause of a given effect, and on the principles according to which they do so (e.g., Shultz, 1982; Shultz & Kestenbaum, 1985; Siegler, 1975; see White, 1995 for review). Thus, one very wide-spread type of task in this context involves presenting children with an outcome O that happens in the presence of two objects (or other events), A and B, and the task is for children to judge whether it is A or B (or both) that causes O (e.g., Bullock, Gellman, & Baillargeon, 1982; Schlottmann, 1999; Shultz & Mendelson, 1975; for a recent variation on this line of research see Gopnik, Sobel, Schulz, & Glymour, 2001; Sobel, Tenenbaum, & Gopnik, 2004). Research using this type of task might broadly be described as investigating the bases on which children judge that a particular object or event has the causal power to produce a certain effect, whereas another object or event lacks that causal power (either because of the type of object or event it is, or because it does not stand in the right sort of relationship with the effect). We will call this type of reasoning causal powers reasoning (the term 'causal power' is used loosely here without implying any commitment to any particular theory of causal learning or knowledge). Much of the debate over the nature and development of causal reasoning has been concerned with establishing the principles and underlying mechanisms that guide causal powers reasoning (Corrigan & Denton, 1996; Gopnik et al., 2004; White, 1995).

As we have said, in typical causal reasoning tasks, participants are provided with some sort of evidence that allows them to infer the causal powers of an object or event. Imagine, for example, a task in which children are provided with types of evidence that would allow them to infer that event A has the causal power to produce an outcome O1 and that event B has the causal power to produce an outcome O2. The research evidence mentioned above would suggest that 4-year-olds should be able under many circumstances to use the evidence that they are provided with in principled ways to infer the causal powers of events A or B (i.e., that they produce O1 and O2, respectively). However, to pass our temporal–causal-reasoning task, it is not enough simply to be aware of the causal powers of the relevant individual events. In our task, there are two events A (the putting of the hairbrush in the blue cupboard) and B (the putting of the hairbrush in the red cupboard) and children have to deduce the *final* outcome of the event sequence—we will call this the FO (The nature of the task is such that in fact the FO will either be O1 or O2.). However, knowing the outcomes of each of the individual events A and B, while trivially easy to infer in our task (e.g., putting the hairbrush in the red cupboard results in the outcome of it being in the red cupboard), is clearly not sufficient to judge the FO correctly. Success on this task depends on bringing to bear an additional awareness that one final outcome, FO1, will obtain if the event sequence was A then B and a different final outcome, FO2, will obtain if the event sequence was B then A. This is what seems to cause 4-year-olds problems, either because they lack the realization that this is the case, or because they fail to use such a realization to guide their reasoning.

Thus, we are making a distinction between conventional causal reasoning tasks that involve making judgments about the causal powers of individual events (i.e., being able to judge the outcomes that they have resulted/result in), and being able to reason about the causal significance of the order in which a series of such events occurred (i.e., being able to judge the *final* outcome that obtains/has obtained after a sequence of these events). Clearly, being able to judge FOs depends upon prior knowledge of individual event outcomes in the event series. That is, it would be impossible to judge the FO if the individual outcomes O1 and O2 for A and B were not known. In fact, a key advantage of our task over that previously used by McCormack and Hoerl (2005) is that in the present task the individual outcomes O1 and O2 of each event A and B are straightforward and do not require pre-training children. In contrast, in McCormack and Hoerl's (2005) study children had to actually take part in a pre-training session in which they learned the novel outcomes O1 and O2 that resulted from A and B (specifically, the different outcomes that resulted from pressing two buttons on a novel piece of apparatus).

As we have said, knowledge of or an ability to deduce individual outcomes (the causal powers of events) is not sufficient to pass temporal–causal-reasoning tasks. Not only are we are making a distinction between two sorts of abilities, we are suggesting that one may emerge later in development than another. However, what might be thought to complicate the picture here is that some of the studies that have assessed children's ability to infer the causal powers of individual events have looked at whether 4-year-olds can use evidence that is specifically temporal in nature. Two types of temporal evidence have been examined. First, research on the temporal priority principle has demonstrated that 4-year-olds' causal judgments usually respect the principle that causes must be temporally prior to effects (Bullock & Gelman, 1979; Bullock et al., 1982). Second, research has examined whether temporal contiguity has an impact on children's causal judgments. Typically, the issue is: if one of the candidate causes happens some time before an outcome, but the other candidate is temporally contiguous with the outcome, do children have a tendency to judge that the latter candidate is the cause of the outcome? A number of studies have found that children do indeed often choose the temporally contiguous event as the one which had the causal power to produce outcome (Schlottmann, 1999; Siegler & Liebert, 1975; though see Shultz, 1982).

Does the fact that 4-year-olds' causal powers judgments often respect such temporally-specified principles mean that in fact they are capable of temporal–causal reasoning as we have described it? We would argue children may be able to exploit such temporal information in choosing which of two candidates A and B caused a particular outcome without having to consider explicitly the temporal relationship between those two candidate causes. In tasks examining use of the temporal priority principle, typically an initial event A occurs, followed by an outcome O, and then a subsequent event B. Use of the principle is demonstrated by a choice of A over B as a cause of O. Potentially, children could choose correctly simply by operating with a default along the lines of: once O has happened, no longer consider any further potential causes. Use of such a default would not actually even involve a consideration of the relative order of A and B. Similarly, use of temporal contiguity as a cue is typically demonstrated in a situation in which an event A happens, followed by a delay, and then event B followed immediately by the effect O. Again, children might choose B because of a strong (albeit potentially illusionary; see Schlottmann, 1999) perceptual effect that leads them to disregard A as a potential cause. Temporal contiguity effects are observed in many animals and at their most basic may be a reflection of the fundamental properties of associative learning mechanisms or systems involved in the storage of perceptual information (White, 1988). Operating with such defaults and heuristics would lead to correct responses without the child actually considering the relative temporal orders of A, B, and O.

It remains to be established if this characterization of the processing underlying 4-year-olds' respect for temporal factors in their judgments of the causal powers of individual events is correct. However, their high levels of success on such tasks is in marked contrast to their poor performance on the tasks used here and in other studies, indicating that the latter tasks do tap a more demanding reasoning ability.

We now turn to briefly considering the ways in which our findings and task analyses relate to research assessing some other important types of cognition that may also be thought to have a temporal component: planning and script use. Some planning tasks (most notably those used by Atance & O'Neill, 2001) require children as young as three to consider a future situation in which their desire differs from their current desire, in line with Tulving's (2004) requirements for a "true" planning task. However, such tasks do not require children to consider the temporal order of a future event sequence, but to consider a single future event (i.e., they are "one step" planning tasks). Tasks such as Atance's trip task simply require that the children know that in order to achieve a certain outcome, they will need to have with them an object with certain causal powers. As such they do not require children to reason about FOs (final outcomes) of an event sequence. In contrast, some other planning tasks do explicitly require children to consider the temporal order in which individual events need to happen to achieve a certain FO (most notably, versions of the Tower of Hanoi/London task; Klahr & Robinson, 1981; Levin et al., 1991; Luciana & Nelson, 1998; Welsh, 1991; Welsh, Pennington, & Groisser, 1991); consistent with our findings, performance on such tasks improves greatly between 3 and 5 years.

Second, research on children's memory for familiar event sequences has amply demonstrated their ability to represent and remember temporal order information (Nelson, 1986; Fivush & Hudson, 1990; reviewed in Nelson, 1996). It might be argued that the ability to exploit a script-like representation of a familiar event sequence requires exactly the sort of temporal–causal reasoning abilities that we have been talking about. Consider an example of a familiar event sequence in which the FO is having the shopping done, where A is selecting goods, and B is paying at the check-out. For the FO to obtain A needs to happen before B. Research on children's use of scripts demonstrates that children can learn the order in which events need to happen in order for FOs such as these to obtain and apply such knowledge appropriately to guide their behavior and expectations. Does this constitute evidence that children can reason about the causal significance of temporal order information? It is important to note that we are not arguing that children cannot learn through experience or observation that the FO of an event sequence AB can be different from the FO of an event sequence BA. That is why our task involves only a single trial, since we acknowledge that given a number of exposures with feedback, such learning could take place. However, exploiting such learning is a matter more akin to recognition than reasoning: it simply involves children recognizing that this is a circumstance in which, for example, FO1 will obtain, rather than insightfully using order information to make a deduction.

In general, what we are suggesting, along with Povinelli et al. (1999), is that children's problems with tasks like the ones used in this study may reflect a failure to fully grasp or utilize in their reasoning the nature of the relationships between events in time. We have argued that these findings are consistent with demonstrations of 4-year-olds' competence on tasks involving other sorts of temporal and causal cognition (causal powers reasoning, planning, and script use). The idea that there is a fundamental change between 3 and 5 years in children's temporal cognition is not a new one. In particular, there have been attempts to link changes in children's emerging competence with temporal relational terms with underlying cognitive changes (Cromer, 1971; Weist, 1986, 1989; Weist, Lyytinen, Wysocka, & Atanassova, 1997; see Friedman, 1978). The studies described here provide additional evidence that 4-year-olds have difficulty with what has been termed temporal–causal reasoning. They also provide further support for the basic distinction between this type of reasoning and more primitive temporal updating abilities. We have suggested, along with Povinelli et al. (1999), that this may reflect a failure to exploit in their reasoning that the temporal order of events has causal significance.

References

Atance, C. M., & O'Neill, D. K. (2001). Planning in 3-year-olds: A reflection of the future self? In C. Moore & K. Lemmon (Eds.), *The self in time. Developmental perspectives* (pp. 121–140). Mahwah, NJ: Erlbaum.

Bauer, P. J., Wenner, J., Dropik, P. L., & Wewerka, S. S. (2000). Parameters of remembering and forgetting in the transition from infancy to early childhood. *Monographs of the Society for Research in Child Development, 65*, 1–204.

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*, 413–426.

Bullock, M., & Gelman, R. (1979). Preschool children's assumptions about cause and effect: Temporal ordering. *Child Development, 50*, 89–96.

Bullock, M., Gellman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–253). New York: Academic Press.

Campbell, J. (1994). *Past, space, and self*. Cambridge, MA: MIT Press.

Carver, L. J., & Bauer, P. J. (2001). The dawning of a past: The emergence of long-term explicit memory in infancy. *Journal of Experimental Psychology: General, 130*, 726–745.

Corrigan, R., & Denton, P. (1996). Causal understanding as a developmental primitive. *Developmental Review, 16*, 162–202.

Cromer, R. F. (1971). The development of the ability to decentre in time. *British Journal of Psychology, 62*, 353–358.

Das Gupta, P., & Bryant, P. (1989). Young children's causal inferences. *Child Development, 60*, 1138–1146.

Fivush, R., & Hudson, J. A. (1990). *Knowing and remembering in young children*. Cambridge, UK: Cambridge University Press.

Fivush, R., & Mandler, J. M. (1985). Developmental changes in the understanding of temporal sequence. *Child Development, 56*, 1437–1446.

Friedman, W. J. (1977). Development of children's understanding of cyclic aspects of time. *Child Development, 48*, 1593–1599.

Friedman, W. J. (1978). Development of time concepts in children. In H. W. Reese & L. P. Lipsitt (Eds.). *Advances in child development and behavior* (pp. 267–298). New York: New York Academic Press.

Friedman, W. J. (1991). The development of children's memory for the time of past events. *Child Development, 62*, 139–155.

Friedman, W. J. (2003a). Arrows of time in early childhood. *Child Development, 74*, 155–167.

Friedman, W. J. (2003b). The development of a differentiated sense of the past and future. *Advances in Child Development and Behavior, 31*, 229–269.

Friedman, W. J., & Brudos, S. L. (1988). On routes and routines: The early development of spatial and temporal representations. *Cognitive Development, 3*, 167–182.

Friedman, W. J., Gardener, A. G., & Zubin, N. R. E. (1995). Children's comparisons of the recency of 2 events from the past year. *Child Development, 66*, 970–983.

Friedman, W. J., & Kemp, S. (1998). The effects of elapsed time and retrieval on young children's judgments about the temporal distances of past events. *Cognitive Development, 13*, 335–367.

Friedman, W. J., & Lyon, T. D. (2005). Development of temporal-reconstructive abilities. *Child Development, 76*, 1202–1216.

Gelman, R., Bullock, M., & Meck, E. (1980). Preschoolers' understanding of simple object transformations. *Child Development, 51*, 691–699.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*, 3–32.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*, 620–629.

Hudson, J. A., Sosa, B., & Shapiro, L. R. (1997). Scripts and plans: The development of preschool children's event knowledge and event planning. In S. L. Friedman & E. K. Scholnick (Eds.). *The developmental psychology of planning: Why, how, and when do we plan?* (pp 77–102). Mahwah, NJ: Erlbaum.

Klahr, D., & Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cognitive Psychology, 13*, 113–148.

Levin, M., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., et al. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology, 7*, 377–395.

Lewkowitz, D. J. (2004). Perception of serial order in infants. *Developmental Science, 7*, 175–184.

Liben, L. S. (2003). Beyond point and shoot: Children's developing understanding of photographs as spatial and expressive representations. *Advances in Child Development and Behavior, 31*, 1–42.

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four-to-eight-year-old children. *Neuropsychologica, 36*, 273–293.

McCormack, T., & Hoerl, C. (2005). Children's reasoning about the causal significance of the temporal order of events. *Developmental Psychology, 41*, 54–63.

Mandel, D. R., Nelson, D. G. K., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development, 11*, 181–196.

Martin, M. (2001). Out of the past: Episodic recall as retained acquaintance. In C. Hoerl & T. McCormack (Eds.). *Time and memory: Philosophical and psychological perspectives* (pp. 257–284). Oxford, UK: Oxford University Press.

Nelson, K. (1986). *Event knowledge: Structure and function in development*. Hillsdale, NJ: Erlbaum.

Nelson, K. (1996). *Language in cognitive development: Emergence of the mediated mind*. Cambridge, UK: Cambridge University Press.

Povinelli, D. J., Landry, A. M., Theall, L. A., Clark, B. R., & Castille, C. M. (1999). Development of young children's understanding that the recent past is causally bound to the present. *Developmental Psychology, 35*, 1426–1439.

Schlottmann, A. (1999). Seeing how it happened and knowing how it works: How children understand the relationship between perceptual causality and underlying mechanism. *Developmental Psychology, 35*, 303–317.

Shultz, T. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development, 47*, 1–51.

Shultz, T., & Kestenbaum, M. (1985). Causal reasoning in children. *Annals of Child Development, 2*, 195–249.

Shultz, T., & Mendelson, R. (1975). The use of covariation as a principle of causal analysis. *Child Development, 46*, 394–399.

Siegler, R. S. (1975). Defining the locus of developmental differences in children's causal reasoning. *Journal of Experimental Child Psychology, 20*, 512–525.

Siegler, R. S., & Liebert, R. M. (1975). Effects of contiguity, regularity, and age on children's causal inferences. *Developmental Psychology, 10*, 574–579.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28*, 303–333.

Tulving, E. (2004). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.). *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3–56). Oxford, UK: Oxford University Press.

Weist, R. M. (1986). Tense and aspect. In P. Fletcher & M. Garman (Eds.) (2nd ed.. *Language acquisition: Studies in first language development*, pp. 356–374). Cambridge, UK: Cambridge University Press.

Weist, R. M. (1989). Time concepts in language and thought: Filling the Piagetian void from two to five years. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 63–118). Amsterdam, The Netherlands: Elsevier.

Weist, R. M., Lyytinen, P., Wysocka, J., & Atanassova, M. (1997). The interaction of language and thought in children's language acquisition: A cross-linguistic study. *Journal of Child Language, 24*, 81–121.

Welsh, M. C. (1991). Rule guided behaviour and self-monitoring on the Tower of Hanoi disk-transfer task. *Cognitive Development, 6*, 59–76.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function in children. *Developmental Neuropsychology, 7*, 131–149.

White, P. A. (1995). *The understanding of causation and the production of action: From infancy to adulthood*. Hove, England: Psychology Press.

White, P. A. (1988). Causal processing: Origins and development. *Psychological Bulletin, 104*, 36–52.